

**INSTITUTE FOR SOCIAL AND ECONOMIC RESEARCH AND POLICY
COLUMBIA UNIVERSITY
WORKING PAPERS**

**‘ACTIVE PATIENTS’ IN RURAL AFRICAN HEALTH CARE:
IMPLICATIONS FOR WELFARE, POLICY AND PRIVATIZATION**

**Kenneth L. Leonard
Department of Economics
Columbia University**

March 2003

ISERP WORKING PAPER 03-03

ISERP
Institute for Social and Economic Research and Policy

Abstract

The 'active patient' is introduced in this paper. She is the same person as the rational peasant that we have known for at least three decades. She is a rational agent seeking health care in an environment characterized by market failures (particularly agency in the supply of medical quality) and imperfect institutional responses to these failures. We show evidence that patients significantly increase their welfare by choosing between various different providers and matching their illnesses to the resources that are available at these different providers. This paper suggests that continuing to view patients as passive participants in the health care market gives way to misleading policy suggestions and may in fact reduce the welfare of patients.

Keywords: Agency in Health Care, rational peasant, rural health care, Africa, asymmetric information.

JEL Classification: I1, O1, O2

The purpose of this paper is to propose a view of patients—particularly the rural poor in Africa—that is different in many aspects from currently accepted views, to show evidence of this view’s validity, and to discuss its implications. The assumptions underlying the ‘active patient’ view have long been accepted when applied to peasants in the context of agricultural production; the ‘rational peasant.’ Our active patient is the same person as the rational peasant but we observe her seeking health care rather than making production or marketing decisions.

We choose the term ‘active patient’ to point out the underlying assumptions in the word *patient*. A standard definition of patient is “an individual awaiting or under medical care and treatment.” However an additional definition is “one that is acted upon.” Sen (1995, p. 11) uses this latter definition in the context of poverty alleviation programs.

“To see [them] as *patients* rather than as *agents* can undermine the exercise ... Not to focus on the fact that they think, choose, act, and respond is to miss something terribly crucial.”

This paper suggests that the same concerns should apply to health care. The behavior of rural patients in Africa can be productively viewed as the behavior of rational individuals maximizing their utility in an environment with pervasive market failures and imperfect institutional responses to these failures. This is in contrast to both a view of passive patients, and to a view of rational patients in a setting with complete markets, or efficient institutional responses to limited market failures.¹

We introduce a model of health seeking behavior that differs from standard models in its view of health, patient’s valuation of health and the environment in which patients seek health. We suggest four important differences. First, we suggest that each illness condition² from which a patient might suffer should be seen as having a unique production function for healthiness. Changes in health status (gains in healthiness) are produced by combinations of factors such as medicine, skill and effort, where the elasticities of healthiness with respect to each of these inputs differs according to the illness. Second, health seeking should be seen as, at least in part, an investment, not a consumption decision. When the patient is sick, increases in health increase the earning potential of the patient. Thus, the budget constraint is endogenous to the health seeking decision (Grossman, 1975). Third, markets for health care suffer from market failures like asymmetric information, or moral hazard (Arrow, 1963, 1986; Dranove and White, 1987; Mooney and Ryan, 1993; Gaynor, 1994). Patients seek medical care precisely because the doctor knows more than they do about health care. Therefore patients will generally not be able to evaluate the quality of care they received from the doctor. They cannot directly purchase the inputs they value, but are instead forced to purchase the inputs that are available through implicit or explicit contracts. Fourth, African patients are seeking this valuable good that suffers from asymmetric information in an environment in which regulatory and institutional guarantees of quality differ greatly from organization to organization. In addition—and this is where our view would differ from a rational–patient–in–complete–markets view—regulation is imperfect and the institutions that govern the delivery of health care are not efficient or optimal in any sense. These institutions are historically based and due to rigidities it is not correct to view them as efficient responses to the needs of any individual patient today.³

A strong implication of the model is that quality is not an absolute feature of practitioners. Some practitioners will provide superior quality for some inputs and others will provide superior quality for other

¹This taxonomy closely follows Stiglitz (1989). Complete markets traditionally assume no market failures. However, with efficient institutions or contracts markets can behave as if they were complete (achieving full information outcomes) even if there are some market failures. Since the implications are the same, we group both views of health care (complete markets and efficient institutional responses to market failures) together, and will refer to this as a complete markets view.

²We use the term *illness condition* as distinct from *diagnosis* for the following reason. We will be assuming that patients use information about their illness condition when they choose a practitioner, and they could not use information about diagnoses because they need to see a practitioner to know their diagnosis. Thus an illnesses condition is the set of information available to a patient before she seeks care.

³See Stiglitz (1989) for a discussion of institutions, institutional history and rigidities.

inputs. Since illness conditions vary in the need for different inputs, the net benefit of visiting a provider will vary according to illness conditions. The same patient might receive the best value of care at one provider when she suffers from one illness, and the best value of care at another provider when she suffers from another illness condition. In Africa, very few rural residents have access to any form of health insurance and they can therefore choose between any of a number of providers for each and every illness condition. We will show that patients do exercise significant choice; the same family might visit a large number of different providers for different conditions over the course of the same year. It is the contention of this paper that these choices can be modeled as rational matches of illness conditions to the resources supplied at various providers.

The model introduced is not very different, in its basic form, from sophisticated models of health seeking behavior (see for example Gertler, Locay and Sanderson, 1987; Dor, Gertler and van der Gaag, 1987; Gertler and Van der Gaag, 1990; Bolduc, Lacroix and Muller, 1996). However, we suggest that patients possess and use much more information about health care than these authors have been willing to assume and that this information plays an important role in the choice of provider. The health seeking behavior of these informed and active patients leads to a very different understanding of issues such as the elasticity of demand for medical care.

On the other hand, our approach is very different than the view of patients which permeates much of the literature and most of the public policy debate on health care delivery. That view is dominated by passive patients and a concern with delivery of services. Health care is seen as a universal human right and therefore the job of all governments is to assure access to affordable standard health services. Though universal access is a beneficial poverty alleviation program, the focus on taking services to patients ignores the contribution that patients can make on their own behalf. It is either assumed that patients do not choose between providers (they go to the closest facility) or that if they do choose facilities it is damaging (or at best unhelpful) to their own health. Most policy makers assume that all use of traditional healers, drug sellers, and self-care are symptoms of ignorance or poverty on the part of consumers. Most of the literature does not model the decision to seek alternative sources of care, the reasons that patients choose facilities that are not the closest facility or the types of information that people possess. We will discuss how this view of patients fails to take advantage of the opportunities that are present, and imposes unnecessary costs on patients.

In this paper, we present empirical evidence of the validity of this view. We show evidence of strong patterns of choice on the basis of illness condition. In one case these patterns are explained by appeal to a theory of the incentives of different providers to provide medical effort and fixed levels of skill. In another case the patterns are compared to objective measures of a variety of inputs into health care (include diagnosis and prescription quality), and the patterns are compared to objective medical opinion as to which illnesses respond more to which inputs. In both cases patients are shown to exhibit patterns of choice that show sophistication in their understanding of both the levels of inputs that are available and the use of those inputs.

The paper is organized as follows. In the next section we develop a stylized model of our view of ‘active patients.’ Section 2 examines the empirical evidence in support of our view. Section 3 discusses the policy implications of this view and Section 4 concludes.

1 An ‘Active Patient’ View of Health Care

Economic theorizing about health care implicitly derives from the “medical model” of Parsons (1951, Chap. 10). Parsons saw the patient as an inadequately informed and *hence passive* recipient of care from an omniscient physician. Observation of patients in consultation with physicians certainly supports this view. The following summaries of conversations with Lebanese women encapsulate the near-universal experience of health-seeking.

Lebanese women's perception of the obstetric care they received was characterized by the feeling of passivity. . . . [W]omen are often very cooperative with the medical team in a hospital setting and don't express their concerns and worries. . . . No woman reported that she had actively challenged any procedure or aspect of care she received. (Kabakian-Khasholian, Campbell, Shediach-Rizkallah and Ghorayeb, 2000, pp. 111–112)

This passivity is a direct result of the asymmetry of information: patients do not know at the time whether what is happening is the right thing or not, and even if they do, it is not at all clear what action they could undertake to improve anything. However, though patients are passive in front of the doctor, they are not passive when they choose doctors.

[W]omen based their choice of provider on their previous experience with the same physician or on what they heard from other women in their social network. (Kabakian-Khasholian et al., 2000, p. 106)

Patients have a choice about who they visit and even though they appear to be—and certainly act—powerless when face to face with a physician, there is no reason for them to be powerless when they choose a physician. Indeed patients are observed to choose between a wide variety of practitioners. The empirical literature on health care in Africa demonstrates quite conclusively that the sick are not passive in their use of the health care system but instead are actively making choices between an array of options. The choices made often are explained by factors unrelated to medicine, such as price and distance, or by static attributes of the facilities, such as availability of drugs or perceived quality (Stock, 1983; Yoder, 1989; Waddington and Enyimayew, 1989; Abel-Smith and Rawal, 1992; Litvack and Bodart, 1993). Another group of studies, however, indicates that choice is being guided by the character or severity of the illness (Mwabu, 1986, 1989; Mwabu and Mwangi, 1986; Mwabu, Ainsworth and Nyamete, 1993; Sauerborn et al., 1989; Bichmann et al., 1991; Leonard, 2003a, 2000; Ndeso-Atanga, 2000). What information do patients have and how do they use it?

1.1 A model of health seeking behavior

We introduce a model of health seeking that relies neither on passive patients, nor on patients who can argue or negotiate with their doctors over the quality of care. We introduce first a model of the value of health care that illuminates two aspects of information possessed by patients. After introducing a model of the value of health, we model the role of health care, allowing us to discuss models of patient choice of health care practitioner.

1.2 The Value of Health

We begin with an individual who has fallen sick.⁴ The patient does not know the disease but she does know the illness condition; the set of symptoms and other characteristics observable to patients. The given level of health stock is H . Health intervention might lead to a change in the level of health, ΔH . We simplify the idea of health intervention by assuming that there are only two possible outcomes; the worst outcome \underline{h} (no

⁴This is an obvious but important beginning. Often health care is seen in a critical light because of its association with the unfortunate event of poor health: "The money used to pay for health care may otherwise have been used for food, agricultural development or education. Payment for health services is thus made at considerable social cost to the family." (Waddington and Enyimayew, 1989, p. 38) The terrible event that has befallen such families is not health care, but illness. In addition, when a patient is healthy (or not suffering from a health crisis) health care is only one of many methods of improving health. When they are sick, however, health care is reasonably seen as a priority intervention.

change in health status, for example) and the best outcome \bar{h} (full recovery, for example).

$$\Delta H \in \{\underline{h}, \bar{h}\} \quad (1)$$

In most cases the best outcome will be the same as returning a patient to her level of health previous to being sick, but there is no need to make such a restriction. Sometimes the best outcome is not as good as the previous state of health (a traumatic accident with permanent injuries, for example) and other times the best outcome can be better than the previous state of health (a heart attack as a result of an unrecognized heart condition that leads the patient to change eating habits or lifestyle).

There are two outcomes and we assume two possible distributions over these outcomes. ϕ^* is the ‘true diagnosis’ distribution and ϕ^\emptyset is the ‘false diagnosis’ distribution.

true diagnosis	ϕ^*	$\text{prob}(\Delta H = \bar{h}) = \rho^*$	$\text{prob}(\Delta H = \underline{h}) = 1 - \rho^*$
false diagnosis	ϕ^\emptyset	$\text{prob}(\Delta H = \bar{h}) = \rho^\emptyset$	$\text{prob}(\Delta H = \underline{h}) = 1 - \rho^\emptyset$

We motivate these distributions as follows; if the patient’s condition is correctly diagnosed and the proper treatment regime is prescribed, understood and followed, the patient will have a probability of full recovery of ρ^* . If the diagnosis is incorrect, the probability of recovery is ρ^\emptyset . Note that it will not generally be true that if everything is done correctly the patient must recover, nor that if nothing is done correctly the patient cannot recover.

The patient does not value health care, only health outcomes. However, in order to increase the probability of being cured the patient might purchase health care. Health care can be represented as a technology that probabilistically spans ϕ^* and ϕ^\emptyset . A better technology (better health care) is one that has a higher probability of choosing the ‘true diagnosis’ distribution than another technology. We represent the technology by Γ ($0 \leq \Gamma \leq 1$) where

$$\Delta H \sim \Gamma \cdot \phi^* + (1 - \Gamma) \cdot \phi^\emptyset \quad (2)$$

The ‘best’ technology ($\Gamma = 1$) has a 100% chance of correct diagnosis and leads to a chance of recovery of ρ^* , and the ‘worst’ technology ($\Gamma = 0$) has a 100% chance of choosing among the incorrect diagnoses and leads to a chance of recovery of ρ^\emptyset . Increases in Γ increase the probability of a patient being cured. The properties of the two binomial distributions (ρ^* and ρ^\emptyset) are given by the illness condition. Γ is a general function of all possible inputs into health care including, but not limited to medicine, laboratory tests, skill, and effort. The distribution of outcomes and their stochastic relationship to diagnoses, is an important source of asymmetric information in health care. Observing a good outcome does not allow patients to infer Γ .

We follow the basic model of Grossman (1975) and consider health as increasing the number of hours available for work and leisure (the investment component) as well as augmenting utility directly (a consumption component.) Utility is a function of health status, income potential minus the costs of health care and the disutility of the effort of the patient. Thus, $U = U(H, I(H) - C, c(p))$. H is health status (or health stock), $I(H)$ is the income potential at health level H , C are monetary costs, p represents non-monetary costs (such as patient effort), and $c(p)$ is the disutility of these non-monetary costs.

Because there are only two possible outcomes, utility is a probabilistic mapping (defined by Γ) of the utility when the patient is cured and the utility when the patient is not cured.

$$EU = \underbrace{\Gamma (\rho^* \bar{U} + (1 - \rho^*) \underline{U})}_{\text{utility under true diagnosis}} + (1 - \Gamma) \underbrace{(\rho^\emptyset \bar{U} + (1 - \rho^\emptyset) \underline{U})}_{\text{utility under false diagnosis}} \quad (3)$$

$$\text{where} \quad \bar{U} = U[\bar{h}, (I(\bar{h}) - C), c(p)] \quad \text{and} \quad \underline{U} = U[\underline{h}, (I(\underline{h}) - C), c(p)]$$

\bar{U} is the utility if the patient is cured and \underline{U} is the utility if the patient is not cured. Here, it helps to assume

a separable utility form such that $U = V[H, I(H)] - C - c(p)$. Although income and total costs are measured in the same units and need not be separated we choose this formulation for the following reasons. The income (or earning potential of the patient) and health level for good outcomes is the same whether the patient sought health care or not; it depends on the outcome, not the process. The part of utility inside the utility operator ($V[H, I(H)]$) depends on the outcome, not on the costs or effort exerted. Thus, in the end, this part of utility will be independent of which provider is chosen whereas the fees paid or the disutility of other patient inputs might depend on the provider chosen. Costs and disutility have a linear relation to utility. For ease of exposition we write $V[\bar{h}, I(\bar{h})]$ as \bar{V} and $V[\underline{h}, I(\underline{h})]$ as \underline{V} .

The expected utility when no medical care is sought ($\Gamma = 0, C = 0$ and $c(p) = 0$) is $\rho^\emptyset \cdot \bar{V} + (1 - \rho^\emptyset) \cdot \underline{V}$. Using the separable utility function, expanding Equation 3 and subtracting the utility when no care is sought the net expected utility is

$$\Delta EU(\Gamma) = \Gamma(\rho^\star - \rho^\emptyset) \cdot (\bar{V} - \underline{V}) - C - c(p)$$

We scale \underline{V} to zero and introduce the term G such that $G = (\rho^\star - \rho^\emptyset)\bar{V}$, and further expand C to include both fees and travel costs $F + T$. G is the maximum expected gain from health care; (ΓG) , is the actual expected gain from health care.

$$\Delta EU(\Gamma) = \Gamma G - F - T - c(p) \quad (4)$$

with

$$\Gamma = \gamma(r_1, r_2, \dots, r_L) \quad (5)$$

r_l are inputs into the health care process (such as medical effort, medical skill, pharmaceuticals, etc.) Thus the net gain in healthiness is the product of the quantity of health care consumed (Γ), the gain in the probability of a good outcome ($\rho^\star - \rho^\emptyset$) and the value of a good outcome ($\bar{V} = V[\bar{h}, I(\bar{h})]$), minus the fees paid (F), the cost of travel to a health care provider (T) and any non-monetary costs of the patient incurred in seeking health care ($c(p)$). For any normal good with these characteristics it remains only to maximize the net gain in utility by setting the marginal contribution of each input equal to its cost.

1.3 Asymmetric Information in the Supply of Health Care Inputs

However, health care is not an ordinary good. Health care suffers from asymmetric information which means that many of the resources that are provided in health care cannot be evaluated by patients. Take medical effort, for example. A patient can know whether or not medical effort is valuable to the cure, but cannot evaluate whether or not medical effort was provided. Thus the patient will not be able to purchase any arbitrary level of any given input. Note that in this model the outcome is binary. Patients will only observe whether or not they were cured and it is difficult to evaluate the quantity or quality of inputs when there are only two possible outcomes.

Even when inputs are observable, (politeness, cleanliness of facilities, availability of key medicines, whether or not there is electricity for the X-ray machine, etc.) institutional rigidities suggest that for many types of providers (government for example) the patient must simply accept the level of input that is provided. The electricity is either on or it is off, patients cannot negotiate with a doctor in a government hospital as to whether or not it will be on.

Thus, patients will not be able to purchase optimal levels of inputs at many providers, and for many inputs they will not be able to purchase optimal levels at any provider. Patients will receive some amount of

various resources at any provider they choose to visit and therefore, we have:

$$\begin{aligned}\Delta EU_{ijk} &= \Gamma_{ijk} G_{ik} - F_{jk} - T_{ij} - c(p) \\ \Gamma_{ijk} &= \gamma_k (r_{1(ijk)}, r_{2(ijk)}, \dots, r_{L(ijk)})\end{aligned}\tag{6}$$

with the terms in Equation 6 defined in Table 1.

Table 1: Definitions of Terms in Equation 6

term	definition
i	Index of individuals
j	Index of providers
k	Index of illness conditions (cough, diarrhea, injury, etc.)
l	Index of inputs (diagnostic effort, quality prescription, etc.)
ΔEU_{ijk}	Net expected utility (for individual i , with illness k at provider j)
Γ_{ijk}	Technology of health care ($0 \leq \Gamma \leq 1$)
γ_k	Production function for technology (Γ), which varies by k , the illness condition
$r_{l(ijk)}$	Quantity of resource or input l provided
G_{ik}	Maximum expected gain from health care ($G = (\rho^* - \rho^0)\bar{V}$)
F_{jk}	Expected fee and drug costs of illness condition j at provider k
T_{ij}	Travel cost of individual i to provider j
$c(p)$	Non-monetary costs in health care such as patient effort

Note that G is independent of the provider chosen, and that the form of the production function for health is also independent of the provider chosen and the patient (though the values of r_l are not). Any characteristics of the patient that are important to the definition of the illness condition (such as age) are characteristics of the illness condition k). There is a separate production function for each illness condition. The set of possible inputs to health is the same, but the usefulness of any input will depend on the illness in question. For some illnesses medicine is more important than diagnosis, for example, for others the opposite will be true.

Definition 1 (Resource elasticity of net healthiness, ε_{lk})

$$\varepsilon_{lk} = \frac{\partial G \Gamma}{\partial r_l} \frac{r_l}{G \Gamma} = \frac{\partial \Gamma}{\partial r_l} \frac{r_l}{\Gamma}\tag{7}$$

The resource elasticity of net healthiness is the percentage change in net healthiness from a 1% change in the supply of resource l .

The elasticity of net healthiness (health outcomes) with respect to resources is a measure of the degree to which outcomes depend on a particular input. This need not be constant, but we suggest that particular illness conditions will have a higher elasticity for particular inputs than other illness conditions. Some resources are more important for some illnesses conditions than for others; ε_{lk} varies by illness condition and resource.

Definition 2 (Optimal resource quantity, r_l^*)

$$r^* \in \underset{\vec{r}}{\operatorname{argmax}} \Delta EU_{ijk} = \gamma(r_1, r_2, \dots, r_l) G_{ik} - F_{jk} - T_{ij} - c(p)\tag{8}$$

r_l^* is an abbreviation of $r_l^*(ik)$, The optimal resource quantity is the optimal level of resource l for each illness condition and patient. r_l^* is an abbreviation of $r_l^*(ik)$.

The optimal level of each resource is a function of the value of health care, the responsiveness of the illness condition to that resource and the unit cost of that resource.

Definition 3 (Actual resource quantity, \hat{r}_l) *The actual resource quantity is the actual amount of each resource provided for patient i , at provider j , for illness condition k . \hat{r}_l is an abbreviation of $\hat{r}_{l(ijk)}$.*

Definition 4 (Estimated resource quantity, \tilde{r}_l) *The estimated resource quantity is the patient's estimate of the quantity of resource l that would be provided for illness condition k at provider j . \tilde{r}_l is an abbreviation of $\tilde{r}_{l(ijk)}$.*

Under the strictest formulation of the rational–patient–in–complete–markets model $r_l^* = \hat{r}_l = \tilde{r}_l$; There are market failures in the delivery of health care, but contracts between providers and patients and competition resolve these issues and insure that the optimal levels of resources are provided.

The purest form of the passive patient model also assumes that $r_l^* = \hat{r}_l$. This is not because patients can observe the use of resources but because the practitioner has a utility function that perfectly matches the patient's utility function over health outcomes and inputs. Doctors, because they are doctors, always insure that their patients receive the best possible care. Note that if doctors are benevolent (or perfect agents) passivity is a rational response for patients.

It is not difficult to find evidence that the traditional view is incorrect. For example, Table 2 show the results of a data collection effort in which doctors were evaluated using vignettes and by observing their consultations of actual patients. The vignette is a case study patient and is similar to an exam. It can be thought of as measuring the knowledge of a clinician. The consultations however, measures practice. There were a series of 26 diagnostic procedures which were required both on the vignette and in the actual consultation. The consultations were observed by other medical doctors who determined what was required. For all diagnostic procedures, 50 percent of doctors knew what was required (yes/yes + yes/no). However, in only half of the consultations observed did these doctors actually do the right thing. Clearly there was an issue of training (all of the doctors should have known what to do in the exam), but there was also an issue of will, since half of those who knew what to do did not. This evidence is not compatible with a view of benevolent physicians.

Table 2: Vignette procedures compared to direct clinician observation

vign/observe pair	obs #	yes/yes %	no/no %	yes/no %	no/yes %
total	3836	25	36	25	15
history taking total	2065	29	31	26	13
physical exam total	1771	19	41	23	17

Each clinician was observed in a vignette (case study patient) as well as in many actual patient consultations. The observations record the number of paired procedures observed in direct consultations. There were 26 vignette/consultation matches in which a clinician is observed in the same situation. yes/yes means the clinician did what was medically required in both the vignette and the actual consultation. yes/no means that he did what was medically required in the vignette but not in the actual consultation. Source: Leonard and Masatu (2002).

Although everyone who works in health care in developing countries would agree that perfect–agent doctors do not exist (or are certainly not common), many continue to view the patient as being passive. As the assumption of benevolent doctors is relaxed, the both the active and the passive patient model would agree that $\hat{r}_l \neq r_l^*$; doctors do not provide optimal effort. However, the passive patient model also assumes that patients have no idea about the production function for health (Γ), particularly ε_{lk} . Without any knowledge of ε_{lk} , there is little value to refined estimates of \tilde{r}_l . Patients may know that $\hat{r}_{lj} > \hat{r}_{lj'}$ (practitioner j provides more of resource l than practitioner j'), but would not be able to differentiate between illnesses for

which this fact was relatively more or less important. Thus patients might form crude opinions of quality at facilities but will be unable to differentiate between illness conditions on the basis of this quality. This view is compatible with models in which patients are choosing between facilities on the basis of overall quality. They prefer high quality facilities if they can afford them, but do not know when they need them.

With active patients, $r_l^* \neq \hat{r}_l$, but \tilde{r}_l (estimated) is at least correlated if not equal to \hat{r}_l (actual). In addition, patients estimate ε_{lk} . It is not clear how they come to possess it, but the empirical evidence will make it clear that they do possess this information.

1.4 How can patients improve their health in this model?

There are two basic strategies available to patients even if they cannot purchase quality directly. First, they can form contracts with health care providers that give the proper incentives to the providers. Second, they can evaluate the incentives that providers face and choose between providers to seek the best *available* quality. We examine the evidence that patients do form contracts when it is possible, and that even when it is not possible they have a sophisticated understanding of the quality that is available at different providers and they choose optimally among providers. This behavior has a significant impact on the health outcomes of the rural poor.

2 Empirical Evidence of ‘Active Patients’

In this section, we examine the contracts that patients form when they are given the opportunity to negotiate directly with practitioners. We then look at the patterns of choice observed in different data sets in Africa. One data set suggests that if you combine the availability of contracts at traditional healers with patient choice you can show that patients choices are optimal responses to the available providers. Another paper examines the link between choice and objectively measured quality and again shows that the patterns of choice suggest that patients are aware of quality and of the differing role of different types of quality in curing their illnesses.

2.1 Traditional Medicine

The general theory of asymmetric information (or agency) has long recognized that one solution to agency is to force the agent (the person performing the unobservable services) to face the loss or gain of his actions. The counterpart in health care is the pay-only-if-cured contract, which has long been recognized as an ideal health care contract (Arrow, 1963; Dranove and White, 1987). If patients paid their doctors only if they were cured, or more if they were cured than if they were not, the economic cost of agency would be reduced.

This ‘ideal’ contract is exactly the method by which traditional healers in Africa do business. Patients in the South West Province of Cameroun pay traditional healers over twice as much if they are cured than if they are not cured (Leonard, 2003a). Healers receive an initial payment and negotiate with the patient over a payment to be made in the future. In all cases, if the treatment did not result in improvement of the condition, the patient paid nothing beyond the initial payment. The pay-only-if-cured contract means that healers have strong incentives to provide quality care even if the patient cannot evaluate or observe this quality. (See Leonard, 2003a, for further discussion of the economics of traditional healers.) The fact that solutions to market failures in the delivery of health care exist in indigenous, traditional institutions is evidence of both an understanding of these market failures on the part of traditional healers, and a willingness to seek these solutions on the part of patients. The existence of this particular contract is, in itself, evidence of active patients.⁵

⁵The irony is that this contract also supports the theory of rational-patient-in-complete-markets. The cultural-based source of

2.2 Self-Selection of Providers by Illness Condition

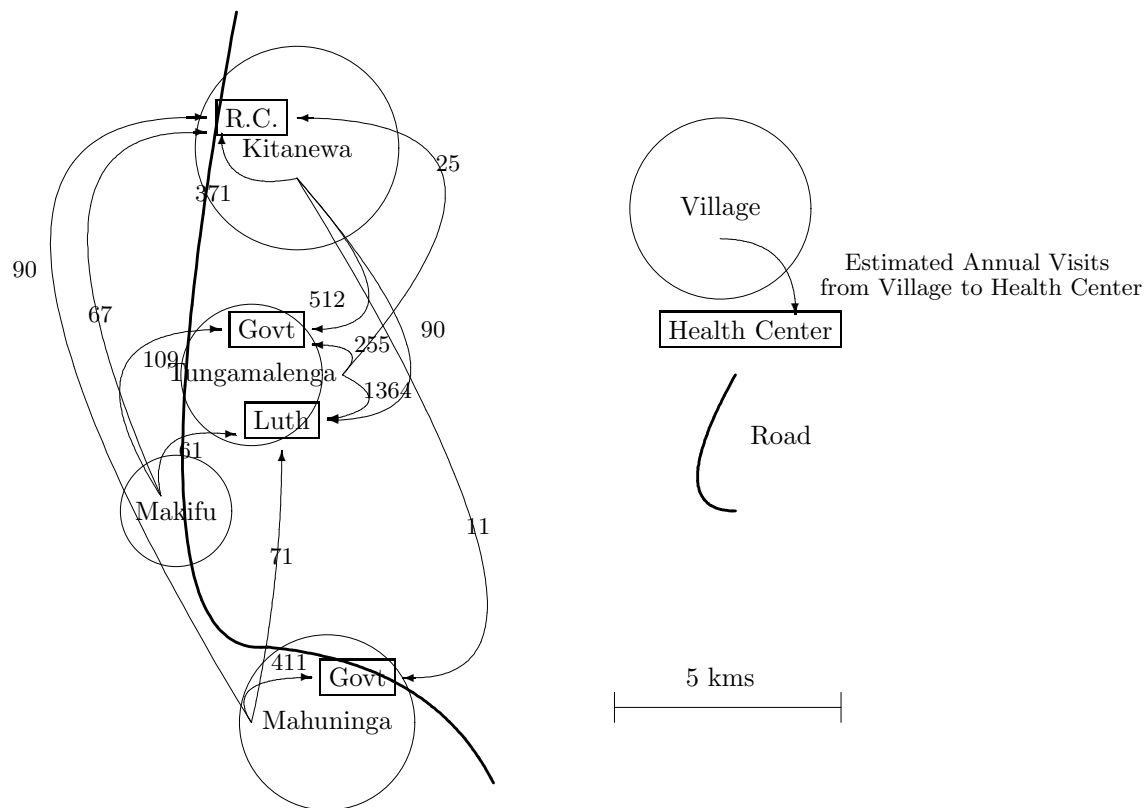


Figure 1: Estimated annual visits to clinics in Idodi Ward, Iringa Rural District, Tanzania

Figure 1 is a stylized representation of the annual visits from a selection of villages to a selection of health facilities in Idodi ward, Iringa Rural District, Tanzania (Leonard, Mliga and Haile Mariam, 2002, p. 1). There are four modern health facilities in the area; two government facilities, one Roman Catholic and one Lutheran facility. There are 4 villages and the radius of each circle representing a village is proportional to the population of that village. The population of Kitaneva (the largest village) is almost 2,400 people. This graph shows two important patterns: patients are bypassing facilities, and patients are bypassing in different directions. Some patients are leaving Tungamalenga to visit a health facility in Kitaneva at the same time that patients Kitaneva are leaving to visit Tungamalenga.

Table 3 shows the outcomes of patients who visited a series of providers (for the first visit) from data collected in the Southwest Province of Cameroun (Leonard, 2003a). In this data, patients can choose between traditional healers, government clinics, government hospitals, mission clinics and mission hospitals. Table 3 reports the outcomes of all illness episodes at different types of practitioners.

Almost 5% of people who visited traditional healers died and healers have the lowest cure rate. At first glance, this supports the commonly-held view of traditional healers as providers of poor quality care who are visited by tradition-bound patients. However, by this measure, the second worst practitioner is the mission hospital. We know that mission hospitals are, beyond question, the highest quality centers. On the other hand, the best cure rates are at government clinics and these are the worst of the modern practitioners.

medicine can be seen as a efficient-institutional response. To find evidence of inefficient, historically-based rigid institutions we need to turn to modern medicine.

Table 3: Episode Outcomes by Provider (Cameroun Household Survey)

Percentage of all visits to a practitioner resulting in given outcome

result	Gov Clinic	Mis Clinic	Gov Hosp	Mis Hosp	Trad Heal	Total
died	1.9	2.7	1.5	3.2	4.6	2.5
cured ^a	79.5	66.9	72.1	59.7	53.0	69.2
sought other care	1.3	2.7	2.2	4.8	4.6	2.6
went home ^b	3.9	4.1	5.2	9.7	4.6	4.9
referred	2.6	3.4	0.7	0.0	3.0	2.1
cont. treatment ^c	10.9	20.3	18.4	22.6	30.3	18.7

Sample of 533

illness episodes from Mbonge subdivision of South West Province of Cameroun. Outcomes are reported by patients; not taken from health facilities.

a Combination of responses “cured” and “well enough”.

b Not cured, not seeking care elsewhere.

c Continuing treatment at the time of the survey.

Source: (Leonard, 2003a).

The findings do not reflect quality, but rather, as we shall see, are the result of self-selection by illness condition. The choice of facility is based, at least in part on the severity of the illness condition.

Table 4 shows the patterns of chief complaint and the first visits to health providers in rural Kenya. Also shown are the clusters of chief complaints generated by a cluster analysis process (Mwabu, 1986). The clusters suggest that visits to providers are being determined (at least in part) by illness conditions (or chief complaints). Note that cluster one and two correspond roughly to chief complaints that lead to a visit to government and mission clinics. Cluster three corresponds to illnesses that lead to visits primarily to traditional healers and mission clinics. Cluster four is visits to mission clinics and government hospitals, cluster five and seven represent visits to mission clinics and pharmacies and cluster six represents visits to government clinics and pharmacies.

(Leonard, 2003a) suggest that patterns of patient choice in Cameroun are evidence of patients who know and understand both ε_{lk} and r_l . The survey in Cameroun was designed to elicit information that patients had before they sought care: all of the symptoms they experienced; the self-declared severity of the disease; the number of days sick before seeking care; and the number of those days in which the patient was bedridden. With the characteristics of the disease plus the age and sex of the individual and information about endemic diseases in the area (but not information on the choice of practitioner or the diagnosis) the paper developed estimates of the elasticities of healthiness (ε_{lk}) with respect to medical effort, medical skill and patient effort (Leonard, 2003a).

The paper follows the same basic model as that shown in Equation 6;

$$\Delta EU_{ijk} = \Gamma_{ijk} G_{ik} - F_{jk} - T_{ij}$$

where $\Gamma = \gamma(\hat{r}_m, \varepsilon_m, \hat{r}_p, \varepsilon_p, \hat{r}_{s(t)}, \varepsilon_{s(t)}, \hat{r}_{s(c)}, \varepsilon_{s(c)}, \hat{r}_{s(h)}, \varepsilon_{s(h)})$

$\hat{r}_{s(t)}$, $\hat{r}_{s(c)}$, and $\hat{r}_{s(h)}$ are the skill provided at traditional healers, clinics (both mission and government) and hospitals (both mission and government) respectively, and are fixed levels. $\varepsilon_{s(t)}$, $\varepsilon_{s(c)}$ and $\varepsilon_{s(h)}$ are the elasticities of outcomes with respect to skills provided at healers, clinics and hospitals respectively. Each facility provides only one level of skill—hospitals always provide at least as much skill as clinics and clinics at least as much as healers—but that skill is more or less useful for different illness conditions.

\hat{r}_m is the quantity of medical effort provided at each provider (for each illness condition) and \hat{r}_p is the quantity of patient effort provided. The quantities of these resources vary with respect to the illness condition. In general $\hat{r}_m = f(\varepsilon_m, \varepsilon_p, \hat{r}_p)$ and $\hat{r}_p = f(\varepsilon_m, \varepsilon_p, \hat{r}_m)$ so we can say that the level of health care

Table 4: Percentage Distribution of first visits by individual illnesses and illness clusters across providers (Rural Kenya)

Cluster	Chief complaint	Govt clinic	Missn clinic	Priv clinic	Govt hosp.	Phmcy or shop	Trad. healer	Self	None
1	Ear	25.0	12.0	12.0	0.0	12.5	0.0	25.0	25.5
1	Eye	40.0	26.7	0.0	6.7	6.7	0.0	6.7	13.3
1	Cough	36.4	27.3	0.0	2.3	22.7	0.0	3.5	13.8
1	Vomiting	40.0	20.0	0.0	0.0	20.0	0.0	0.0	20.0
1	Backache	36.1	27.8	8.3	0.0	11.1	11.1	0.0	5.5
1	Abdomen	41.5	16.9	4.6	1.5	24.6	3.1	6.2	1.5
1	Rib pain	30.0	20.0	10.0	10.0	30.0	0.0	0.0	0.0
1	Diarrhea	25.0	35.0	0.0	0.0	20.0	5.0	15.0	0.0
	Mean %	34.3	23.3	4.4	2.6	18.5	2.4	7.0	8.3
2	Wounds	52.6	15.8	0.0	0.0	0.0	10.5	15.8	5.2
2	Fainting	66.7	33.3	0.0	0.0	0.0	0.0	0.0	0.0
	Mean %	59.7	24.6	0.0	0.0	0.0	5.3	7.9	2.6
3	Asthma	20.0	40.0	0.0	0.0	0.0	40.0	0.0	0.0
3	Bodypain	23.5	17.7	11.8	0.0	5.9	29.4	5.9	5.9
3	Joint pain	20.0	20.0	0.0	6.7	6.7	20.0	13.3	6.7
3	Other	28.6	14.3	14.3	14.3	0.0	28.6	0.0	0.0
	Mean %	23.0	23.0	6.6	5.3	3.0	29.5	4.8	3.2
4	Malaria	0.0	50.0	0.0	37.7	0.0	0.0	7.1	7.1
4	Leprosy	0.0	60.0	0.0	20.0	0.0	20.0	0.0	0.0
	Mean %	0.0	55.5	0.0	28.9	0.0	0.0	3.5	3.5
5	Swelling	20.0	60.0	0.0	0.0	20.0	0.0	0.0	0.0
5	Heart	10.0	40.0	10.0	20.0	20.0	0.0	0.0	0.0
	Mean %	15.0	50.0	5.0	10.0	20.0	0.0	0.0	0.0
6	Headache	21.9	9.5	2.9	0.9	47.6	3.8	4.8	8.6
6	Fever	17.7	8.8	0.0	0.0	58.8	0.0	5.9	8.8
	Mean %	19.8	9.2	1.5	0.5	53.2	1.9	5.4	8.7
7	Tuberculosis	0.0	66.7	0.0	0.0	33.3	0.0	0.0	0.0

Source: Mwabu (1986)

provided is a function of the elasticities. Thus

$$\Gamma = \gamma(\varepsilon_m, \varepsilon_p, \hat{r}_{s(t)}, \varepsilon_{s(t)}, \hat{r}_{s(c)}, \varepsilon_{s(c)}, \hat{r}_{s(h)}, \varepsilon_{s(h)})$$

G is estimated as a linear function of patient characteristics and the severity of the illness condition.

We develop a series of reduced form hypotheses based on the incentives of each provider to provide effort. Clearly, we expect skill to be a significant determinant of choice and that the coefficients on the elasticities of outcomes with respect to all levels of skill ($\varepsilon_{s(t)}, \varepsilon_{s(c)}, \varepsilon_{s(h)}$) will be significant and positive. In addition, analysis of the incentives at providers leads to the following testable hypotheses; mission facilities (both clinics and hospitals) are more likely to be chosen when the elasticity of outcomes with respect to medical effort is high (since these facilities provide uniformly high levels of medical effort but also have lower costs); government facilities are more likely to be chosen when the elasticity with respect to medical effort is low (they provide little effort) and traditional healers are more likely to be chosen when the elasticity with respect to both medical and patient effort are high (traditional healers provide high levels of medical effort, but particularly high levels when patients also provide high levels of effort). All of these hypotheses are born out in the reduced form analysis contained in Leonard (2003a) and are support of the hypothesis that patients are aware of both resource provision levels and the benefit of these resources.

Leonard (2000) uses the same data set and estimates a maximum likelihood structural estimation of the parameters of the contracts available at each provider. Figure 2 is a two dimensional representation of illness conditions according to the elasticity of healthiness with respect to medical and patient effort. Shown on the graph is the maximum likelihood choice of provider holding other features of patients constant. The prediction that patients prefer traditional healers when they suffer from conditions that are very responsive to both medical and patient effort is shown in this graph.

These data cannot be directly compared to the data from Mwabu (1986) because the various providers face different incentives and the definition of chief complaint is not the same as the definition of illness condition. However, the incentives facing traditional healers are the same in both data sets⁶ and it is not difficult to see how asthma (a condition likely to lead to a visit to a traditional healer according to our analysis and Figure 2) body pain and joint pain can be characterized as being illnesses that require effort of both the doctor and the patient as well as a degree of cooperation between them. It is not difficult to imagine these patterns being based on the same type of choice mechanism, again supporting the hypotheses of information and choice in patients.

2.3 Choice and Objective Measures of Quality

In Iringa rural district of Tanzania, Leonard et al. (2002) examines patient choice as well as objective measures of various qualities of care available at a series of clinics owned by the government and nongovernmental organizations (NGOs). All of the facilities are clinics and are similar in terms of capacity but differ markedly on other measures of quality. Quality scores were gathered by other doctors and nurses who visited each facility and include the following:

CONSULT: The average quality of consultations from the point of view of an observing clinician. The score of each facility is the average of the fixed effects for each of its clinicians.

PRESCRIP: This score reflects the appropriateness of the prescriptions given to patients. The ‘perfect’ prescription involves *only* the necessary drugs and *all* the necessary drugs.

INJECT: The percentage of non-infant prescriptions for malaria that were given by injection. This is a per-facility average intended to reflect a proclivity to injections in general. This variable is only indirectly

⁶Personal communication with the author.

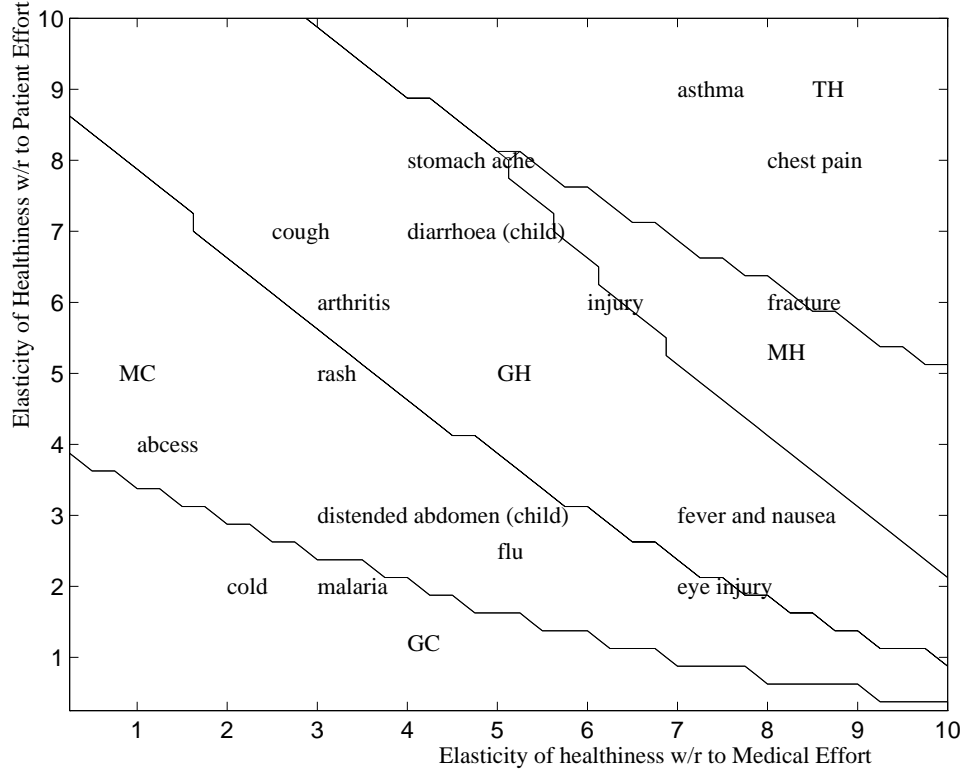


Figure 2: Pattern of Predicted First Visits to Selected Providers according to responsiveness to medical and patient effort (Rural Cameroun)

Source: Leonard (2000). TH traditional healer, MH mission hospital, MC mission clinic, GH government hospital, GC government clinic

measured (as opposed to the first two variables) by looking at outpatient registers and we therefore have less confidence in it.

N DRUGS: Number of drugs prescribed. Unlike PRESCRIP, this is not necessarily in conflict with the health of patients, but inflicts unnecessary costs. This is a per-facility average. This variable is indirectly measured in the same manner as INJECT.

Each facility in the sample has a unique mixture of these scores and we can observe patients choosing between facilities. Leonard et al. (2002) develop a mixed multinomial/conditional logit regression where the dependent variable is the choice between the three nearest government, and two nearest NGO clinics. Included are a series of organization and illness interaction effects to control for expected costs (fees and the cost of drugs) as well as some patient characteristics and the travel distance. Last, we include illness condition/resource interaction terms with CONSULT, INJECT, PRESCRIP, and N DRUGS as resources.

This specification can be represented as the following modification of Equation 6;

$$\Delta EU_{ijk} = \Gamma_{ijk} + G_{ik} - F_{jk} - T_{ij}$$

$$\Gamma = \sum_{l=1}^L \sum_{k=1}^K r_{lj} \cdot \varepsilon_{lk}$$

Each resource (indexed by l) is assumed fixed at each provider (thus r_{lj}) and there is a unique elasticity of outcomes for each illness condition and each resource (thus ε_{lk}). We have visited each provider and can

estimate the levels of r_{lj} and we use a mixed multinomial/conditional logit specification to determine patient estimates of ε_{lk} . We have data on F_{jk} and T_{ij} but we do not have good information on G (only age and gender).

The paper shows that each resource has a significant impact for the average illness; on average, patients seek good quality consultation and prescriptions and avoid facilities that over-use injections and over-prescribe medication. However, patients with different illnesses have different preferences for these different resources: ε_{lk} vary significantly between illness conditions. Patients suffering from particular illness conditions exhibit patterns of choice that differ from the average condition; they are more likely to visit facilities with certain characteristics, for example.

Table 5 shows the patterns that are predicted by a medical analysis of the illnesses in question: medical estimates of ε_{lk} . Shown are the overall preferences; medical opinion states that patients should prefer facilities that have good consultation quality (CONSULT) and good prescription (PRESCRIP) and avoid facilities that over-use injections (INJECT) or use too many drugs (N DRUGS). When medical opinion states that particular illness conditions have different needs for resources from the average condition, this is noted with a +, or -. For example, according to medical opinion, the care-giver of an infant with dysentery has less need of high quality consultation than the average patient and more need of a good prescription. In other words, the condition is not difficult to diagnose, but requires care in the use of medicine.

Also shown in Table 5 is whether or not the medical opinion is supported by the observed patterns. ✓ indicates that medical opinion and the actions of patients agree. ✗ indicates that medical opinion and the actions of patients disagree. Where neither ✓ or ✗ is indicated, agreement or disagreement cannot be statistically determined. + ✓(− ✓) means that, according to medical opinion, patients should seek this resource more (less) than for the average condition and that patient behavior confirms medical opinion.

With the exception of consultation quality for infants (CONSULT) all of the averages agree. On average patients, agree with medical opinion in terms of the usefulness of resources for deviations from the average condition. For the two scores in which the study had the most confidence, CONSULT and PRESCRIP, the agreements are quite good. Even for the other two scores there are more agreements than disagreements.

This empirical investigation is an attempt to estimate ε_{lk} after having directly measured \hat{r}_l and to see whether the patients estimates of resource elasticities and resource levels are similar to medical opinion. The results show that, not only do patients behave as if they know something about the levels of inputs provided, they also know when these resources are of use and for which illness conditions. The can neither accept nor reject perfectly informed patients, but can clearly reject passive patients.

3 Implications for Research and Policy

There is strong evidence that patients have and are using information in their choice of health care facilities. Certainly more work is required in order to gain confidence in this view. However, such work is only useful if the policy implications of this view of patients are significantly different than policy implications that would be obtained under prevailing views.

Demand for health care There is much debate on the price elasticity of demand for health care in developing countries (see, for a good summary Gertler and Van der Gaag, 1990). We suggest, first that such studies will obtain misleading results and second that the more important question is the price elasticity of actual health care resource provision. Each individual study might obtain an estimate of price elasticity of health care, but since patients do not value health care (but rather being cured), what does this demand curve represent? If the quality remains constant, then the study will obtain an elasticity estimate for that level of quality, but unless quality is the same in every study in every country, the results of these studies cannot be compared. Each one measures a response to a particular level of quality.

Table 5: Medical opinion for illness condition/quality matches compared to actual patterns of visits in rural Tanzania

	CONSULT		PRESCRIP		INJECT		N DRUGS	
Infant (Avg)	+		+	✓	-	✓	-	✓
malaria	-		-		+			✓
upper rti		✓		✗	+			✓
pneumonia	-	✓	+		+	✓	-	✗
cough	+	✓	+	✓	-	✓		✗
diarrhea	+		-		-	✓	+	
eye problem		✓	-	✓	-		+	
injury	-		-	✓	+	✗	+	✓
dysentery	-		+		-		-	✗
scabies	-		-		-	✓		✓
worms	-			✓	-	✗		✗
Non-Infant (Avg)	+	✓	+	✓	-	✓	-	✓
malaria		✓	-			✓	-	✓
upper rti	-	✓		✗	-		+	
injury	-	✓	-		+			✓
pneumonia	-		+		+	✗	+	
cough	+	✓	+			✗	+	✓
dysentery	-	✗	+	✓	-	✓		✓
diarrhea	-	✓	-	✓	-	✓		✓
worms	-	✗	-		-	✓		✗
severe abd pain	+		+	✓	+			✗
eye problem	+			✓	-			✓
Pelvic Infl. Dis.	+		+	✗		✗	+	✓
skin inf.		✓	+		-	✗		
sexually transmitted disease	+			✓	+	✗	+	✗
scabies	-		-		-		-	✗
asthma	+	✓		✓	+			✓
correct matches	11		10		9		12	
incorrect matches	2		3		7		8	

✓: Prediction and estimate agree statistically.

✗: Prediction and estimate disagree statistically.

Source: (Leonard et al., 2002). See definitions of terms on 12.

Far more interesting are studies such as Litvack and Bodart (1993), in which quality was improved simultaneously with an increase in price. In that study the demand for some socio-economic groups (including the poor) increased. This study is evidence of the view of active patients we have proposed in this paper. If institutions were rational, or if patients could bargain over inputs, patients should have been able to increase quality without any government reform (since they were willing to pay for it). The paper also suggests what should be the more important question: what is the price elasticity of health care quality, or of particular inputs?

Leonard (2003b) estimates a model in which there are imperfect markets for medical effort. Given the parameters estimated, the paper can simulate the utility that patients would achieve if they could bargain directly with practitioners over the delivery of quality. Shown in Table 6 is the estimated utility (derived from a model of patient choice) of patients under the current system as well as the estimated utility if

Table 6: Estimates of Utility from health care, with and without Agency

utility	trad heal	gov cln	gov hos	mis cln	mis hos	sample
With Agency						
average for all episodes	5.2	12.4	13.7	14.0	6.6	
average for predicted visit	9.6	7.1	22.2	14.6	41.0	15.0
Without Agency; Full Information						
average for all episodes	44.8	55.1	33.0	50.7	22.6	

All units 1,000 CFA (approx 2 USD). Average for all episodes is the average utility for each illness episode if the patient visited that facility. Average for predicted visit is the average utility for patients who choose to visit that facility (utility is weighted by the probability that the patient would choose that facility). Source: (Leonard, 2003b).

patients could bargain over the provision of inputs. The first row is the utility that all patients would achieve if they visited any particular provider. The second row represents the average utility for those illnesses that we predict will lead to visits to that practitioner. The second row reflects the choice of patients. Note that the sample average utility of 15,000 CFA is higher than the average if all patients visited government clinics (12,400 CFA). This is despite the fact that patients have to travel further and pay more money to visit government hospitals, mission clinics or mission hospitals. They improve their utility by wisely exercising choice.

The third row shows the utility that patients would obtain at any given provider in the absence of asymmetric information. This is the utility that would be obtained if patients could negotiate with providers over the supply of inputs. The utility is not weighted by choice (since there would be less use for choice) and should therefore be compared to the utility in the first row. Note that this is the utility after compensating providers for the additional inputs provided. Average utility is greatly increased. Clearly, increasing quality even with higher prices will benefit patients.

Microeconomic foundations for government intervention in the delivery of outpatient services The traditional justification for government intervention in the delivery of outpatient curative services⁷ is that health care is a merit good, or that health care is a universal human right. However, this argument is not sufficient. Food is a merit good by this argument, but most people would agree that governments should not operate restaurants. Patients value health care and are willing to pay for it, why should governments deliver heavily subsidized health care? If they cannot afford health care, why not give them the money spent on the current system and let them purchase what they think is most beneficial to them?

Figure 3 is a suggestion for another possible justification for government intervention in the delivery of curative services in rural areas. The fixed costs of opening a clinic in such areas is very high, and the incomes of people are low. It is possible that delivering services in such an area would result in losses for all providers even if they had a monopoly in a given rural area. The ideal government intervention in such a situation is to have the government subsidize a private practitioner and then regulate the services that he provides. This is common practice for veterinary care in Scandinavian countries. Veterinarians are given stipends, houses and vehicles to operate private practices in under-served rural areas. However, prices and practices are regulated.

African governments are not capable of effective regulation, and this is one of the reasons that government provision of care has been so poor. However, it is possible that even without any regulation, giving a fixed subsidy to a practitioner who operates in a rural area could increase the welfare of the population. The cost of the fixed subsidy ($ABCP_m$) is less than the net gain to the population from paying even the

⁷Vaccinations, non-patient related preventive care and health education present strong justifications for government intervention that we do not discuss here. In addition, hospitalization should be covered by an insurance program for which there are good justifications for government intervention.

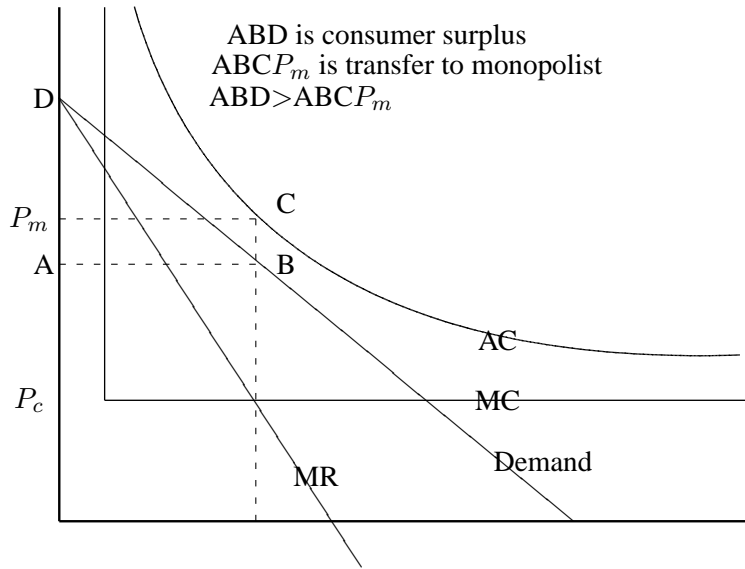


Figure 3: Natural monopoly in the delivery of rural services and the potential welfare gains of subsidized monopoly

monopolist price (ABD). They would pay the monopolist price for health care (P_m) not the competitive price (P_c), but would receive services that were not previously delivered. In this case, giving patients the value of the subsidy ($ABCP_m$) would not result in health care services being provided and the government program is superior to a pure poverty alleviation program.

Whether or not this would work and whether or not the services provided would be of sufficient quality, depends on the price elasticity of quality, something for which we have insufficient measures. However, there is the strong possibility that this unregulated subsidy in the private market would be better than the existing system.

Quality Improvement From the point of view of an institution that provides health services, the fact that patients can recognize and respond to quality means that reforms intended to increase quality can have a double benefit. First, all visitors to a center which increases its quality will benefit and secondly, new patients will be attracted to this center. Patients who had previously incurred significant travel cost to avoid the center will benefit from increased quality and decreased avoidance costs. Travel costs are a transaction cost in the context of health seeking behavior. They reflect a willingness to purchase health care that is transferred to other sectors of the economy (bus drivers, for example) and not retained in the health sector. When patients bypass facilities that are equipped to handle their cases (but do not deliver the services due to asymmetric information) this is a significant loss of value. (For discussion see Leonard, ed (2000, pp. xxxii–xxxiii) and Haile Mariam (2000), as well as Litvack and Bodart (1993)). Increases in quality that lead to less bypassing are a unilateral reduction in waste.

Institutional Reform The assumption of this paper—that patients are forced to choose between imperfect institutions—suggests the possibility for institutional reform. The major failing of the institutions that we have discussed in this paper, and especially of government institutions, is that they do not address the incentives of clinicians to provide quality care. Programs to increase quality tend to focus on training,

equipment and pharmaceutical supply. We do not question the importance of these inputs, but if clinicians have poor incentives to use these inputs, such programs are unlikely to have much impact.

Among the institutions from which patients can choose there are examples of institutions that, though not perfect, are better than others; nongovernmental organizations (NGOs). In Cameroun and Tanzania these are exclusively church-operated institutions. Although they deliver medicine in much the same way as the government services, they pay more attention to incentives and the quality of care they deliver is better. Mliga (2000) discusses some of the organizational practices that lead to this improvement in quality. More generally Leonard (2002) discusses mechanisms for extending the success of these organizations by broadening their impact and using their skills to improve government services (see also Gilson et al., 1997). The success or failure of these mechanisms depends crucially on the reaction of patients to the new institutional forms. Active patients is a theory that allows us to develop a framework for analyzing these reactions.

Health Insurance The lack of comprehensive health insurance or even of private insurance is a severe problem in Africa. There is no question that it imposes a significant burden on patients. However, health insurance usually ties patients down to one particular institution. Although the insurance could be beneficial overall, it will restrict the ability of patients to actively choose between practitioners for every illness condition. Thus, with health insurance, there will be an increase in sub-optimal matching of illness conditions to providers. Clearly, if patients freely choose health insurance they are better off with health insurance. We should be cautious of situations in which patients have to be forced to accept health insurance (which is frequently the case) and some thought should be given to mechanisms which improve access without suppressing choice.

Private Practice One of the more important and ongoing experiments in health care is that of privatization. The analysis of privatization in the light of active patients is an important illumination of the power of this theory. A rule-bound patient should be a discouraging prospect for any provider thinking of opening a private practice. If patients do not understand quality and are not willing to pay for services that are more expensive than the currently available government services, there is not much room for a private practitioner to make a living. On the other hand, the private practitioner should be enthusiastic about selling goods and services to the rational patient if markets are complete. If the quality of the services he sells are of sufficient quality, he should be able to attract a large number of fee paying customers.

Private practice has yet to take off in Africa and has certainly not done so in the rural areas. In those parts of the developing world where privatization has succeeded it specializes in selling certain types of diagnostic tests (ultrasound exams, for example) and pharmaceutical supplies. Private practitioners do not generally offer a broad range of services (Bennett, McPake and Mills, eds, 1997). If patients are so poor that they cannot afford services even if they are offered at marginal cost, the failure of private practice does not imply that the active patient model is not correct; active patients can be too poor to purchase private health care. However, we observe relatively expensive NGO providers succeeding in the same conditions in which private practitioners do not succeed.

Does the failure of private practice imply either that patients are too poor, or that patients are rule-bound rather than rational? Not necessarily so, because it is possible for inefficient institutions to survive and provide services in an environment characterized by market failures where private practice cannot survive. Both the traditional healer (who is a type of private practitioner) and government and NGO services have access to particular mechanisms that allow them to provide and receive compensation for their services. Traditional healers use the pay-only-if-cured contract and government and NGO providers use hierarchical supervision to insure quality. It is important to recognize that the fee for service contract with hierarchical supervision

is not a balanced-budget contract.⁸ A private practitioner cannot use a pay-only-if-cured contract because outcomes are difficult to observe and even more difficult to verify.⁹ Nor, however, can a private practitioner use hierarchical supervision (or any form of broken budget contract) because he is not part of a hierarchy. Thus private practitioners might fail where other types of providers would not.

This suggests two possible future lines of investigation on the prospects of privatization. Private practitioners have traditionally been seen as the practice of an individual practitioner, but this does not have to be the case. If practitioners can become part of a larger hierarchy they can potentially gain from the types of contracts used in NGOs. This does not mean that they have to become employees of a large organization, they might become active members in a medical association, or part of a network based on mutual referrals. Such mechanisms do not currently exist in Africa, but this does not mean that they should not or could not.

The second line of inquiry is about the possibility of reputation overcoming the market failures that arise from asymmetric information. The data examined in Section 2.3 suggested that patients act as if they know the level of quality available at any given center. This does not imply complete markets because patients might know quality without being able to impact it. However, it does suggest that patients might be learning about quality by observing outcomes over a long period of time. If this is so, then private practitioners might succeed when patients have had the opportunity to learn about quality. We do not know how long it takes for patients to evaluate the reputation at a given facility, or more importantly how long it takes for patients to start responding to their beliefs about quality. If patients are rule-bound, then by definition, they never learn about quality. However, if they do learn, but slowly, a private practitioner cannot hope to gain sufficient return on his investment in a time frame that makes it worth his effort. Knowing how patients learn is essential to answering this question.

4 Conclusions

We do not suggest that there is a fundamentally new idea contained in our view of active patients. The model of such an actor has existed for at least three decades. Nor is there anything in the application of the ‘rational peasant’ to health care economics that is controversial or difficult to accept. None-the-less the two existing models of patient behavior—the rule-bound patient and the rational-patient-in-complete-markets view—ask different questions and see the data differently. In health policy circles in developing countries it is the rule-bound patient theory that holds sway and the policy implications from this view are very different from the policy implications that come from active patients.

The energy that patients exert in the search for health care is enormous. Patients learn about quality, how it applies to different illnesses and about symptoms and their relationships to illness. They travel to seek care, even bypassing closer facilities. When patients are not cured they travel even further in search of cures. These energies should not be wasted in attempts to circumvent poorly designed institutions. In an ideal policy setting these energies would be harnessed and turned into productive forces in the health care system.

By viewing patients as rule-bound, these energies are ignored. When policy analysts assume that patients are rule-bound they often work against them. When fees are raised at government clinics, ministry officials observe that patients are no longer attending and conclude that they are no longer seeking care. The fees are dropped, and the reforms that could have been paid for with the extra revenue are shelved. In fact patients were simply turning to other sources of health care, and if the reforms had been implemented they might have returned and welfare would have increased (Leonard and Leonard, forthcoming).

⁸A balanced budget requires all payments from the patient to be equal to payments to the practitioner; a institution can pay the practitioner less without collecting less or more without collecting more (see Hölmstrom, 1982; Clarke, 1971; Groves, 1973).

⁹Outcomes are non-contractible (McGuire, 2000, p. 499). Traditional healers have a unique solution to the non-contractibility issue; patient believe that if they lie about outcomes they will be punished by the spirits that affected their cure (Leonard, 2003a).

We have shown some empirical evidence which, though limited, should be sufficient to warrant a closer look at the view of active patients. Two different data sets and empirical investigations have shown sophisticated patients making informed decisions between health care providers that increase their welfare. It is truly a mistake at this point to continue to view patients as passive participants in their own health care.

References

Online documents are available at www.columbia.edu/~kl206/research.html

- Abel-Smith, B. and P. Rawal**, “Can the Poor Afford ‘Free’ Health Services? A Case Study of Tanzania,” *Health Policy and Planning*, 1992, 7 (4), 329–41.
- Arrow, K. J.**, “Uncertainty and the Welfare Economics of Medical Care,” *American Economic Review*, 1963, 53 (5), 941–973.
- , “Agency and the Market,” in K. J. Arrow and M. D. Intriligator, eds., *Handbook of Mathematical Economics*, Vol. 3, Amsterdam: Elsevier Science Publishers, 1986, pp. 1183–1195.
- Bennett, S., B. McPake, and A. Mills, eds**, *Private Health Providers in Developing Countries: Serving the Public Interest?*, London and New Jersey: Zed Books, 1997.
- Bichmann, W. et al.**, “District Health Systems: Users’ Preferences for Services in Benin,” *Health Policy and Planning*, 1991, 6 (4), 361–70.
- Bolduc, Denis, Guy Lacroix, and Christophe Muller**, “A comparison of medical providers in rural Benin: a comparison of choice based models,” *Journal of Health Economics*, 1996, 15 (4), 477–498.
- Clarke, E.**, “Multipart pricing of public goods,” *Public Choice*, 1971, 8, 19–33.
- Dor, Avi, Paul Gertler, and Jacques van der Gaag**, “Non-Price Rationing and the Choice of Medical Care Providers in Rural Cote D’Ivoire,” *Journal of Health Economics*, December 1987, 6 (4), 291–305.
- Dranove, D. and W. D. White**, “Agency and the Organization of Health Care Delivery,” *Inquiry*, 1987, 24, 405–415.
- Gaynor, M.**, “Issues in the Industrial Organization of the Market for Physician Services,” *Journal of Economic Management and Strategy*, 1994, 3 (1), 211–255.
- Gertler, P. and J. Van der Gaag**, *The Willingness to Pay for Medical Care: Evidence from two developing countries*, Baltimore, Maryland: Johns Hopkins University Press, 1990. Published for The World Bank.
- , **L. Locay, and W. Sanderson**, “Are user fees Regressive? The welfare implications of health care financing proposals in Peru,” *Journal of Econometrics*, 1987, 36, 67–68.
- Gilson, L. et al.**, “Should African Governments contract out clinical health services to church providers?” In Bennett et al., eds (1997) chapter 17.
- Grossman, Michael**, “On the Concept of Health Capital and the Demand for Health,” *Journal of Political Economy*, 1975, 80, 223–255.
- Groves, T.**, “Incentives in Teams,” *Econometrica*, 1973, 41, 617–631.
- Haile Mariam, Damen**, “Traditional Insurance Mechanisms and the Choice of Health Care Providers in Ethiopia.” In Leonard, ed (2000) chapter 2.
- Hölmstrom, Bengt**, “Moral Hazard in Teams,” *Bell Journal of Economics*, 1982, 13, 324–40.
- Jack, W.**, *Principals of Health Economics for Developing Countries*, Washington, D.C.: The World Bank, 1999.
- Kabakian-Khasholian, T., O. Campbell, M. Shediak-Rizkallaha, and F. Ghorayebc**, “Women’s experiences of maternity care: satisfaction or passivity?,” *Social Science and Medicine*, 2000, 51 (1), 103–113.
- Leonard, David K., ed.**, *Africa’s Changing Markets for Human and Animal Health Services*, London: Macmillan, 2000.
- Leonard, Kenneth L.**, “Incentives and Rural Health Care Delivery.” In Leonard, ed (2000) chapter 4.

- , “When States and Markets Fail: Asymmetric Information and the Role of NGOs in African Health Care,” *International Review of Law and Economics*, 2002, 22 (1), 61–80.
- , “African Traditional Healers and Outcome–Contingent Contracts in Health Care,” *Journal of Development Economics*, 2003, 71 (1). (available on-line).
- , “Imperfect Agency in Health Care: An Empirical Investigation in Rural Cameroun,” mimeo (available on-line), Columbia University 2003.
- **and David K. Leonard**, “The Political Economy of Improving Health Care for the Poor in Rural Africa: Institutional Solutions to the Principal-Agent Problem,” *Journal of Development Studies*, forthcoming.
- **and Melkiory C. Masatu**, “Using Vignettes to evaluate quality in outpatient services: Tanzania,” mimeo (available on-line), Columbia University 2002.
- , **Gilbert Mliga, and Damen Haile Mariam**, “Bypassing Health Facilities in Tanzania: Revealed Preferences for Observable and Unobservable Quality,” *Journal of African Economies*, December 2002, 11 (4).
- Litvack, J. I. and C. Bodart**, “User fees plus quality equals improved access to health care: results of a field experiment in Cameroun,” *Social Science and Medicine*, 1993, 37 (3), 369–383.
- McGuire, T.**, “Physician Agency,” in A.J. Culyer and J.P. Newhouse, eds., *Handbook of Health Economics*, Amsterdam: Elsevier Science Publishers, 2000, pp. 461–536.
- Mliga, Gilbert R.**, “Decentralization and the Quality of Health Care.” In Leonard, ed (2000) chapter 8.
- Mooney, Gavin and Mandy Ryan**, “Agency in Health Care: getting beyond first principles,” *Journal of Health Economics*, 1993, 12, 125–135.
- Mwabu, Germano M.**, “Health Care Decisions at the Household Level: Results of a Rural Health Survey in Kenya,” *Social Science and Medicine*, 1986, 22 (3), 315–19.
- , “Nonmonetary factors in household choice of medical facilities,” *Economic Development and Cultural Change*, 1989, 37 (2), 383–92.
- **and Wilfred M. Mwangi**, “Health Care Financing In Kenya: A simulation of welfare effects of user fees,” *Social Science and Medicine*, 1986, 22 (7), 763–767.
- , **M. Ainsworth, and A. Nyamete**, “Quality of Medical Care and Choice of Medical Treatment in Kenya: An empirical analysis,” *Journal of Human Resources*, 1993, 28 (4), 838–862.
- Ndeso-Atanga, Sylvester**, “Health Care Quality and the Choice of Care Providers in Rural Cameroon.” In Leonard, ed (2000) chapter 5.
- Parsons, T.**, *The Social System*, Glencoe, Ill: The Free Press of Glencoe, 1951.
- Sauerborn, R. et al.**, “Low Utilization of Community Health Workers: Results from a Household Interview Survey in Burkina Faso,” *Social Science and Medicine*, 1989, 29 (10), 1163–74.
- Sen, Amartya**, “The Political Economy of Targeting,” in Dominique van de Walle and Kimberly Nead, eds., *Public Spending and the Poor*, Baltimore, MD: Johns Hopkins University Press, 1995.
- Stiglitz, J. E.**, “Rational Peasants, Efficient Institutions, and a Theory of Rural Organization: Methodological Remarks for Development Economics,” in P. Bardhan, ed., *The Economic Theory of Agrarian Institutions*, Oxford: Oxford University Press, 1989, chapter 2, pp. 18–29.
- Stock, Robert**, “Distance and the Utilization of Health Facilities in Rural Nigeria,” *Social Science and Medicine*, 1983, 17 (9), 563–570.
- Waddington, C.J. and K.A. Enyimayew**, “A price to pay: The impact of user charges in Ashanti-Akim district, Ghana,” *International Journal of Health Planning and Management*, 1989, 4, 17–47.

Yoder, R.A., “Are People Willing and Able to Pay for Health Services?,” *Social Science and Medicine*, 1989, 29 (1), 35–42.

Recent ISERP Working Papers

03-01: “The Plasticity of Participation: Evidence From a Participatory Governance Experiment,” Shubham Chaudhuri, Economics, Columbia University, and Patrick Heller, Sociology, Brown University

03-02: “Factional Politics and Credit Networks in Revolutionary Vermont,” Henning Hillmann, Sociology, Columbia University

03-03: “ ‘Active Patients’ in Rural African Health Care: Implications for Welfare, Policy and Privatization,” Kenneth L. Leonard, Economics, Columbia University

02-01 “Alternative Models of Dynamics in Binary Time-Series-Cross-Section Models: The Example of State Failure,” Nathaniel Beck, Political Science, UC San Diego, David Epstein, Political Science, Columbia University, Simon Jackman, Political Science, Stanford University and Sharyn O’Halloran, Political Science, Columbia University

02-02 “Substitutability Cross-Stream Between Oriented Markets: Conventions in the Wine Sector of France,” Harrison White, Sociology, Columbia University

02-03 “Link, Search, Interact: The Co-Evolution of NGOs and Interactive Technology,” Jonathan Bach, Center on Organizational Innovation, Columbia University and David Stark, Center on Organizational Innovation, Columbia University

02-04 “Chains of Affection: The Structure of Adolescent Romantic and Sexual Networks,” Peter Bearman, Institute for Social and Economic Research and Policy, Columbia University, James Moody, Sociology, Ohio State, Katherine Stovel, Sociology, University of Washington

02-05 “Permanently Beta: Responsive Organization in the Internet Era,” Gina Neff, Center on Organizational Innovation (COI), Columbia University, and David Stark, Center on Organizational Innovation (COI), Columbia University

02-06 “Negotiating the End of Transition: A Network Approach to Political Discourse Dynamics, Hungary 1997,” Balázs Vedres, Columbia University, Péter Csigó, Ecole des Hautes Etudes en Sciences Sociales

02-07 “The Influence of Women and Racial Minorities Under Panel Decision-Making on the U.S. Court of Appeals,” Sean Farhang, Political Science, Columbia University, Gregory Wawro, Political Science, Columbia University

02-08 “The Role of Effort Advantage In Consumer Response to Loyalty Programs: The Idiosyncratic Fit Heuristic,” Ran Kivetz, Graduate School of Business, Columbia University, and Itamar Simonson, Graduate School of Business, Stanford University

For copies of ISERP Working Papers
visit http://www.iserp.columbia.edu/initiatives/working_papers/paper_program.html,
write to iserp@columbia.edu or call 212-854-3081

EDITORIAL BOARD

Karen Barkey, Sociology
Peter Bearman, Sociology/ISERP
Alan Brinkley, History
Charles Cameron, Political Science
Alessandra Casella, Economics
Ester Fuchs, Political Science/SIPA
John Huber, Political Science
Ira Katznelson, Political Science/History
Herbert Klein, History
Mary Clare Lennon, Public Health
Mahmood Mamdani, Anthropology
Marianthi Markatou, Statistics
William McAllister, ISERP
Kathryn Neckerman, ISERP
Richard Nelson, Business/SIPA
Elliot Sclar, Architecture, Planning and
Preservation/SIPA
Seymour Spilerman, Sociology
Charles Tilly, Sociology
Harrison White, Sociology

ADMINISTRATION

Peter Bearman, Director
Kathryn Neckerman, Associate Director
Leslie Wright, Assistant Director

Institute for Social and Economic
Research and Policy
Columbia University
International Affairs Building
420 West 118 Street, 8th Floor
Mail Code 3355
New York, NY 10027
telephone: 212-854-3081
facsimile: 212-854-8925
e-mail: iserp@columbia.edu
URL: <http://www.iserp.columbia.edu>